

217 #10



OIPE

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:39

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw

ENTERED

```

3 <110> APPLICANT: Lanes, Olav
4      Willasen, Nils Peder
5      Guddal, Per Henrik
6      Gjellesvik, Dag Rune
8 <120> TITLE OF INVENTION: Cod uracil-DNA glycosylase, gene coding therefore,
9      recombinant DNA containing said gene or operative parts
10     thereof, a method for preparing said protein and the
11     use of said protein or said operative pa
13 <130> FILE REFERENCE: U013209-3
15 <140> CURRENT APPLICATION NUMBER: 09/758,017A
16 <141> CURRENT FILING DATE: 2001-01-10
18 <150> PRIOR APPLICATION NUMBER: 2000 5428
19 <151> PRIOR FILING DATE: 2000-10-27
21 <150> PRIOR APPLICATION NUMBER: 2000 0163
22 <151> PRIOR FILING DATE: 2000-01-12
24 <160> NUMBER OF SEQ ID NOS: 19
26 <170> SOFTWARE: PatentIn Ver. 2.0
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1283
30 <212> TYPE: DNA
31 <213> ORGANISM: Gadus morhua
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (18)..(920)
37 <400> SEQUENCE: 1
38 gacatccgct tgcaaat atg ttg ttc aag tta ggg tta tgc caa aga tgc      50
39           Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys
40           1           5           10
42 ata tca tca aat cgg gtg tta cca ggt tta cta att ccc caa act tta      98
43 Ile Ser Ser Asn Arg Val Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu
44           15           20           25
46 tgt ttt tct aaa tta atg aag ata acg ccg aag aaa ctg agg tcc tca      146
47 Cys Phe Ser Lys Leu Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser
48           30           35           40
50 aat gtg gaa caa aag acg tca tcg cca cag ctt tca gtg gag cag ctg      194
51 Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu
52           45           50           55
54 gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag att aga gca      242
55 Glu Arg Met Ala Lys Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala
56           60           65           70           75
58 aaa gca acg cct gca ggt ttc gga gag act tgg aga aga gag ctg gct      290
59 Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala
60           80           85           90

```

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:39

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw

```

62 gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc ttt gta gct 338
63 Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala
64          95          100          105
66 gat gag agg agc cgt cac acc gtc tac cca ccg gct gat caa gtg tac 386
67 Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr
68          110          115          120
70 agt tcg aca gag atg tgt gac att caa gat gtg aaa gta gtg att cta 434
71 Ser Ser Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu
72          125          130          135
74 ggc cag gac cct tac cac ggt ccc aac caa gca cat gga ctc tgt ttc 482
75 Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly Leu Cys Phe
76 140          145          150          155
78 agt gtg caa aag cca gtt ccc cct ccc ccc agt ctc gtg aac ata tac 530
79 Ser Val Gln Lys Pro Val Pro Pro Pro Pro Ser Leu Val Asn Ile Tyr
80          160          165          170
82 aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct gga cat gga 578
83 Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro Gly His Gly
84          175          180          185
86 gat cta agc gga tgg gca aaa caa ggg gtg ctg ctg ctt aac gcg gtg 626
87 Asp Leu Ser Gly Trp Ala Lys Gln Gly Val Leu Leu Leu Asn Ala Val
88          190          195          200
90 ctg acc gtg cgg gcc cat cag gcc aac tcc cac aag gac aga ggc tgg 674
91 Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp Arg Gly Trp
92          205          210          215
94 gag acc ttc acc gac gct gtg atc aag tgg ctg agc gtc aac cgg gaa 722
95 Glu Thr Phe Thr Asp Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu
96 220          225          230          235
98 gga gtc gtt ttc ctg ttg tgg ggc tca tac gcc cat aag aag gga gcg 770
99 Gly Val Val Phe Leu Trp Gly Ser Tyr Ala His Lys Lys Gly Ala
100          240          245          250
102 acc atc gac agg aaa cgt cac cat gtc ttg caa gct gtt cat cca tct 818
103 Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Val His Pro Ser
104          255          260          265
106 cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac ttc tcc aag 866
107 Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys
108          270          275          280
110 gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata aac tgg aga 914
111 Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg
112          285          290          295
114 gca ctc taactcttta tgctgcctta tactgttaac tgttttaaga tgaacatcac 970
115 Ala Leu
116 300
118 actatatattt ctacagctttt tccaagttca aaccaatcta taagctttca tttgtctttt 1030
120 ggaatgatgc tgcttttggt cggtttttaga tacttaaaac actttaccac tctgccatgt 1090
122 tgactcatgt tcagtcataa taactttcac aacttgaaca aaaatgttat tttataattg 1150
124 attatattct gtacattaaa gattgttttt tcccaggct gtttcatagg tactaggata 1210
126 ttaaactgtt attaacctat tttccatgat gtcaactgct taagttttta tgcagaaata 1270
128 aattatatat tta 1283
131 <210> SEQ ID NO: 2

```

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:39

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw

```

132 <211> LENGTH: 301
133 <212> TYPE: PRT
134 <213> ORGANISM: Gadus morhua
136 <400> SEQUENCE: 2
137 Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys Ile Ser Ser Asn Arg
138   1           5           10           15
140 Val Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu Cys Phe Ser Lys Leu
141           20           25           30
143 Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys
144           35           40           45
146 Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu Glu Arg Met Ala Lys
147           50           55           60
149 Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala Lys Ala Thr Pro Ala
150   65           70           75           80
152 Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala Ala Glu Phe Glu Lys
153           85           90           95
155 Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala Asp Glu Arg Ser Arg
156           100          105          110
158 His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr Ser Ser Thr Glu Met
159           115          120          125
161 Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu Gly Gln Asp Pro Tyr
162           130          135          140
164 His Gly Pro Asn Gln Ala His Gly Leu Cys Phe Ser Val Gln Lys Pro
165 145          150          155          160
167 Val Pro Pro Pro Pro Ser Leu Val Asn Ile Tyr Lys Glu Leu Cys Thr
168           165          170          175
170 Asp Ile Asp Gly Phe Lys His Pro Gly His Gly Asp Leu Ser Gly Trp
171           180          185          190
173 Ala Lys Gln Gly Val Leu Leu Leu Asn Ala Val Leu Thr Val Arg Ala
174           195          200          205
176 His Gln Ala Asn Ser His Lys Asp Arg Gly Trp Glu Thr Phe Thr Asp
177           210          215          220
179 Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu Gly Val Val Phe Leu
180 225          230          235          240
182 Leu Trp Gly Ser Tyr Ala His Lys Lys Gly Ala Thr Ile Asp Arg Lys
183           245          250          255
185 Arg His His Val Leu Gln Ala Val His Pro Ser Pro Leu Ser Ala His
186           260          265          270
188 Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys Ala Asn Gly Leu Leu
189           275          280          285
191 Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg Ala Leu
192           290          295          300
195 <210> SEQ ID NO: 3
196 <211> LENGTH: 1355
197 <212> TYPE: DNA
198 <213> ORGANISM: Gadus morhua
200 <220> FEATURE:
201 <221> NAME/KEY: CDS
202 <222> LOCATION: (90)..(992)

```

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:39

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw

204 <400> SEQUENCE: 3

```

205 gatggttttag gaggatagta ctttgacact ggtagcgaa ggggaaaacg gagttattgt 60
207 gcatatcggt ttagccctac gtttaaaaa atg att ggt caa cag cat atc aac 113
208                               Met Ile Gly Gln Gln His Ile Asn
209                               1                               5
211 tct ttc ttc tca cca gtt tca aaa aag aga gtt tca aag gaa tta ggt 161
212 Ser Phe Phe Ser Pro Val Ser Lys Lys Arg Val Ser Lys Glu Leu Gly
213 10                               15                               20
215 aaa acc gaa aag cat gcc gaa gaa gtt cag ata acg ccg aag aaa ctg 209
216 Lys Thr Glu Lys His Ala Glu Glu Val Gln Ile Thr Pro Lys Lys Leu
217 25                               30                               35                               40
219 agg tcc tca aat gtg gaa caa aag acg tca tgc cca cag ctt tca gtg 257
220 Arg Ser Ser Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val
221                               45                               50                               55
223 gag cag ctg gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag 305
224 Glu Gln Leu Glu Arg Met Ala Lys Asn Lys Lys Ala Ala Leu Asp Lys
225 60                               65                               70
227 att aga gca aaa gca acg cct gca ggt ttc gga gag act tgg aga aga 353
228 Ile Arg Ala Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg
229 75                               80                               85
231 gag ctg gct gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc 401
232 Glu Leu Ala Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser
233 90                               95                               100
235 ttt gta gct gat gag agg agc cgt cac acc gtc tac cca ccg gct gat 449
236 Phe Val Ala Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp
237 105                               110                               115                               120
239 caa gtg tac agt tgg aca gag atg tgt gac att caa gat gtg aaa gta 497
240 Gln Val Tyr Ser Trp Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val
241 125                               130                               135
243 gtg att cta ggc cag gac cct tac cac ggt ccc aac caa gca cat gga 545
244 Val Ile Leu Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly
245 140                               145                               150
247 ctc tgt ttc agt gtg caa aag cca gtt ccc cct ccc ccc agt ctc gtg 593
248 Leu Cys Phe Ser Val Gln Lys Pro Val Pro Pro Pro Pro Ser Leu Val
249 155                               160                               165
251 aac ata tac aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct 641
252 Asn Ile Tyr Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro
253 170                               175                               180
255 gga cat gga gat cta agc gga tgg gca aac aag ggg tgc tgc tgc tta 689
256 Gly His Gly Asp Leu Ser Gly Trp Ala Asn Lys Gly Cys Cys Cys Leu
257 185                               190                               195                               200
259 acg cgc tgc ctg acc gtg cgg gcc cat cag gcc aac tcc cac aag gac 737
260 Thr Arg Cys Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp
261 205                               210                               215
263 aga ggc tgg gag acc tcc acc gac gct gtg atc aag tgg ctg agc gtc 785
264 Arg Gly Trp Glu Thr Ser Thr Asp Ala Val Ile Lys Trp Leu Ser Val
265 220                               225                               230
267 aac cgg gaa gga gtg gtt ttc ctg ttc tgg ggc tca tac gcc cat aag 833
268 Asn Arg Glu Gly Val Val Phe Leu Phe Trp Gly Ser Tyr Ala His Lys

```

RAW SEQUENCE LISTING

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:39

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw

```

269          235          240          245
271 aag gga gcg acc atc gac agg aaa cgt cac cat gtc ttg caa gct ctt 881
272 Lys Gly Ala Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Leu
273          250          255          260
275 cat cca tct cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac 929
276 His Pro Ser Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His
277 265          270          275          280
279 ttc tcc aag gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata 977
280 Phe Ser Lys Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile
281          285          290          295
283 aac tgg aga gca ctc taactcttita tgctgcctta tactgttaac tgttttaaga 1032
284 Asn Trp Arg Ala Leu
285          300
287 tgaacatcac actatatTTTT ctacagcttt tccaagttca aaccaatcta taagctttca 1092
289 tttgtctttt ggaatgatgc tgcttttggt cggttttaga tacttaaaac actttaccac 1152
291 tctgccatgt tgactcatgt tcagtcaata taactttcac aacttgaaca aaaatggtat 1212
293 tttataattg attatatattct gtacattaaa gattgttttt ttcccaggct gtttcatagg 1272
295 tactaggata ttaaactgtt attaacctat tttccatgat gtcaactgct taagttttta 1332
297 tgcagaaata aattatatat tta 1355
300 <210> SEQ ID NO: 4
301 <211> LENGTH: 301
302 <212> TYPE: PRT
303 <213> ORGANISM: Gadus morhua
305 <400> SEQUENCE: 4
306 Met Ile Gly Gln Gln His Ile Asn Ser Phe Phe Ser Pro Val Ser Lys
307 1 5 10 15
309 Lys Arg Val Ser Lys Glu Leu Gly Lys Thr Glu Lys His Ala Glu Glu
310 20 25 30
312 Val Gln Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys
313 35 40 45
315 Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu Glu Arg Met Ala Lys
316 50 55 60
318 Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala Lys Ala Thr Pro Ala
319 65 70 75 80
321 Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala Ala Glu Phe Glu Lys
322 85 90 95
324 Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala Asp Glu Arg Ser Arg
325 100 105 110
327 His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr Ser Trp Thr Glu Met
328 115 120 125
330 Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu Gly Gln Asp Pro Tyr
331 130 135 140
333 His Gly Pro Asn Gln Ala His Gly Leu Cys Phe Ser Val Gln Lys Pro
334 145 150 155 160
336 Val Pro Pro Pro Pro Ser Leu Val Asn Ile Tyr Lys Glu Leu Cys Thr
337 165 170 175
339 Asp Ile Asp Gly Phe Lys His Pro Gly His Gly Asp Leu Ser Gly Trp
340 180 185 190
342 Ala Asn Lys Gly Cys Cys Cys Leu Thr Arg Cys Leu Thr Val Arg Ala

```

VERIFICATION SUMMARY

DATE: 02/14/2002

PATENT APPLICATION: US/09/758,017A

TIME: 10:10:40

Input Set : A:\U0132093.app

Output Set: N:\CRF3\02142002\I758017A.raw